

Pengenalan Pola Suara Manusia Berekstensi File WAV Menggunakan Metode *Fast Fourier Transform* dan *Bayes*

Nurlia Apriani¹, Rita Wiryasaputra², Lastri Widya Astuti³

^{1,2,3} Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Indo Global Mandiri, Jln. Jend. Sudirman No. 629 KM 4, Palembang, 30129, Indonesia

E-mail: ¹nurlia_apriani@yahoo.com, ²ritasaputra@uigm.ac.id, ³lastriwidya@uigm.ac.id

Abstrak

Suara manusia merupakan suatu gelombang bunyi yang sangat unik. Hal itu dikarenakan tiap manusia mempunyai jenis gelombang bunyi yang berbeda. Perbedaan yang sangat mendasar dalam suara manusia ialah tinggi rendahnya suara manusia tersebut yang berhubungan dengan sinyal dari gelombang suara manusia. Tujuan penelitian ini ialah mengetahui hasil akurasi dari metode Fast Fourier Transform dan Bayes dalam pengenalan pola suara. Metode Fast Fourier Transform digunakan untuk Ekstraksi ciri sedangkan metode Bayes digunakan untuk menghitung nilai probabilitas suara antara data latih dan data uji, selanjutnya Bayes digunakan untuk menentukan hasil pengenalan dari beberapa data template latih yang telah disimpan sebelumnya. Penelitian ini dibuat menggunakan Matlab R2016a, dengan cara pencocokan pola suara manusia yang telah dibuat sebelumnya atau disebut dengan data latih dengan pola suara yang baru atau disebut data uji. Beberapa pengujian telah dilakukan antara lain pengujian sistem dengan suara orang yang ada dalam database dan pengujian dengan suara orang yang tidak ada didalam database. Hasil pengujian untuk suara yang ada didalam database sebesar 96% untuk laki-laki pertama dan 76% untuk perempuan pertama. Sedangkan pengujian untuk suara yang tidak ada didalam database sebesar 46% untuk laki-laki kedua dan 50% untuk perempuan kedua.

Kata kunci—Bayes, Database, Fast Fourier Transform (FFT), Matlab R2016a, Pengenalan Suara.

Abstract

The human voice is a very unique sound wave. That's because every human being has a different kind of sound wave. The fundamental difference in human voice is high the low the sound level associated with the signal from sound waves. The purpose of this research is to know the accuracy result from Fast Fourier Transform and Bayes method in pattern recognition. The Fast Fourier Transform method is used for feature extraction and Bayes method is used to calculate the sound probability value between the train data and test data, then Bayes Method is used to determine the result of the introduction of some previously stored train data. This research was made using Matlab R2016a, by matching the pattern of human sound that has been made before or called train data with new sound pattern or called test data. Testing is done on voice in the database and the voice is not in the database. Test results for voice in the database were 96% for first men and 76% for first women. While testing for voice is not in the database is 46% for second men and 50% for second women.

Keywords—Bayes, Database, Fast Fourier Transform (FFT), Matlab R2016a, Voice Recognition.

1. PENDAHULUAN

Pengenalan pola merupakan bidang dalam pembelajaran mesin dan dapat diartikan sebagai tindakan mengambil data mentah dan bertindak berdasarkan klasifikasi data. Salah satu aplikasinya adalah pengenalan suara (*speech recognition*), yang merupakan teknologi yang menggunakan peralatan dengan sumber masuknya adalah suara, seperti mikropon untuk menginterpretasikan suara manusia[1]. *Speech recognition* adalah proses yang dilakukan komputer untuk mengenalikata yang diucapkan oleh seseorang tanpa mempedulikan identitas orang terkait [2]. Suara manusia merupakan suatu gelombang bunyi yang sangat unik, hal itu dikarenakan tiap manusia mempunyai jenis gelombang bunyi yang berbeda-beda, tidak mungkin ada yang sangat sama bunyinya. Meskipun suara manusia berbeda-beda tetapi sebagai pendengar tetap dapat mengenali apa yang diucapkan. Perbedaan yang sangat mendasar dalam suara manusia ialah tinggi rendahnya suara manusia tersebut yang berhubungan dengan sinyal dari gelombang suara manusia. Suara merupakan sesuatu yang dapat didengar dan memiliki ciri sinyal tertentu, sedangkan ucapan merupakan suara yang terdiri dari kata-kata yang diucapkan. Pengucapan dalam sebuah bahasa terbagi atas konsonan (k) dan vokal (v). Pengenalan suara atau ucapan merupakan salah satu upaya agar suara dapat dikenali atau diidentifikasi sehingga dapat dimanfaatkan [3]. Dalam Kamus Besar Bahasa Indonesia (KBBI) huruf vokal sendiri terdiri atas 5 (lima) huruf yaitu, a, e, i, u dan o, sedangkan huruf konsonan terdiri dari atas 21 (dua satu) huruf yaitu, b, c, d, f, g, h, j, k, l, m, n, p, q, r, s, t, v, w, x, y dan z. Pengenalan pola suara yang akan dibuat ini ialah dapat mengubah suara yang telah tersimpan dalam bentuk format *wav* menjadi *text* yang dikenali. *WAV (WAVE-form)* singkatan dari istilah dalam Bahasa Inggris *waveform audio format* merupakan standar format berkas audio yang dikembangkan oleh *Microsoft* dan *IBM*. *Wav* adalah format utama untuk menyimpan data audio mentah pada *Windows*. Pemilihan format *wav* yang digunakan dalam sistem pengenalan pola suara ini dikarenakan format data didukung penuh di dalam matlab, serta merupakan bentuk file asli sehingga bersifat murni dan memiliki *noise* yang rendah.

Agar dapat mengenali suara yang diucapkan maka digunakanlah suatu metode untuk ekstraksi ciri dan pengenalan. Metode yang digunakan dalam penelitian ini yaitu metode *Fast Fourier Transform* sebagai ekstraksi ciri dan metode *Bayes* sebagai pengenalan. Algoritma *Fast Fourier Transform(FFT)* adalah suatu algoritma untuk menghitung *Discrete Fourier Transform (DFT)* yang digunakan untuk menghitung spektrum frekuensi sinyal. Serta *FFT* merupakan prosedur penghitungan *DFT* yang efisien sehingga akan mempercepat proses penghitungan *DFT* yang secara substansial dapat lebih menghemat waktu dari pada metoda yang konvensional. Sedangkan metode *Bayes* merupakan metode yang baik didalam mesin pembelajaran berdasarkan data training, dengan menggunakan probabilitas bersyarat sebagai dasarnya[4].

Beberapa penelitian yang menjadi referensi dalam penelitian ini, diantaranya: Referensi yang pertama mengacu pada penelitian yang membahas tentang perancangan sistem pengenalan pola suara menggunakan metode *Piecewise Dynamic Time Warping (PDTW)*. Metode ekstraksi yang digunakan dalam penelitian ini adalah *Linear Predictive Coding (LPC)* dan pengenalan menggunakan metode *Piecewise Dynamic Time Warping (PDTW)*. Hasil penelitian didapatkan *Word Error Rate (WER)* sebesar 45% untuk 40 data uji dari 4 kata[1]. Selanjutnya referensi kedua mengacu pada penelitian yang membahas tentang penerapan metode ekstraksi ciri *FFT* dan Metode Klasifikasi *K-Nearest Neighbor (KNN)* pada pencarian chord, menunjukan hasil akurasi sebesar 82,35% menggunakan 7 (tujuh) ketetanggaan terdekat[5]. Referensi ketiga membahas tentang bagaimana mengetahui jenis-jenis tanaman herbal dengan cara mengidentifikasi bentuk citra daun dari tanaman tersebut. Metode *Naïve Bayes Classifier* merupakan Metode *Bayesian Learning* yang paling cepat dan sederhana. Sedangkan metode *KNN* dapat melakukan klasifikasi dengan cepat berdasarkan jarak terdekat diantara objek data. Berdasarkan hasil uji coba yang dilakukan, penggunaan metode *Naïve Bayes Classifier*

didapatkan nilai akurasi sebesar 75%, sedangkan dengan menggunakan metode *K-Nearest Neighbor* didapatkan nilai akurasi sebesar 70,83%[6]. Kemudian referensi keempat membahas tentang bagaimana menentukan ketebalan tempurung dan daging buah kelapa sawit menggunakan teorema *Bayes*, pada penelitian ini didapat tingkat keakuratan mencapai 84,333% [7].

1.1. Rumusan Masalah

Bagaimana mengimplementasikan metode *Fast Fourier Transform (FFT)* dan *Bayes* dalam pengenalan pola suara manusia berekstensi file *wav*.

1.2 Tujuan dan Manfaat

Tujuan dari penelitian ini adalah membangun sistem pengenalan pola suara berekstensi file *wav* yang berpola kata *kvkv* dan *kvkkv* serta mengetahui hasil akurasi dari metode *Fast Fourier Transform* dan *Bayes* dalam pengenalan pola suara.

Sedangkan manfaat yang diharapkan dari penelitian ini yaitu, *voice recognition* dapat digunakan untuk mengenali suara yang menghasilkan *text*, dalam penelitian ini kata-kata yang digunakan ialah yang memiliki pola kata *kvkv* dan *kvkkv* yang telah ditentukan. Serta memberi informasi mengenai hasil akurasi dari penelitian pengenalan pola suara menggunakan metode *Fast Fourier Transform* dan *Bayes*.

2. METODE PENELITIAN

Adapun metode yang digunakan pada penelitian ini adalah sebagai berikut:

2.1 Metodologi Penelitian

Teknik pengumpulan data yang digunakan dalam penelitian ialah:

- Studi Literatur.
Teknik pengumpulan data dengan cara mengumpulkan jurnal, *paper* dan bacaan-bacaan yang kemudian dirangkum sebagai bahan acuan penyusunan laporan. Penulis mendapatkan referensi melalui sumber internet maupun data dari perpustakaan. Tujuannya adalah untuk memperkuat permasalahan serta sebagai dasar teori.
- Observasi.
Teknik pengumpulan data dengan merekam suara yang akan di simpan dan dijadikan sebagai *template* data latih. Perekaman suara dilakukan pada 4 (empat) orang responden, yaitu 2 (dua) orang laki-laki dan 2 (dua) orang perempuan. Sedangkan untuk kata yang diucapkan ialah bola, dadu, fakta, foto, guru, harta, lampu, mata, salju dan waktu.

2.2 Model Pembangunan Perangkat Lunak

Model pembangunan perangkat lunak yang digunakan dalam sistem pengenalan pola suara ini adalah model *waterfall* yang menyediakan pendekatan alur hidup perangkat lunak secara sekuensial atau terurut. Adapun tahap-tahap dalam pembangunan perangkat lunak dengan model *waterfall* atau sekuensial linier menurut [8] adalah sebagai berikut:

- Analisis kebutuhan perangkat lunak
Tahap analisis pada sistem pengenalan pola suara yang akan dibangun ini yaitu menganalisa data apa saja yang diperlukan, seperti kebutuhan data suara untuk dikenali yang memiliki kriteria pola kata *kvkv* dan *kvkkv*, dan menentukan perangkat lunak dan perangkat keras yang sesuai dengan sistem yang akan dibuat.

- **Desain**
Tahap desain pada sistem yang akan dibangun yaitu membuat rancangan untuk pengenalan pola suara secara keseluruhan.
- **Pengkodean**
Desain harus ditranslasikan ke dalam program perangkat lunak. Hasil dari tahap ini adalah program komputer sesuai dengan desain yang telah dibuat pada tahap sebelumnya. Pengkodean perangkat lunak merupakan proses penulisan bahasa program dimana pada penelitian ini yaitu *Matlab* sehingga perangkat lunak tersebut dapat dijalankan oleh mesin.
- **Pengujian**
Peneliti menggunakan pengujian data *training* dan data *testing* pada tahap pengujian. Hal ini dilakukan untuk meminimalisir kesalahan (*error*) dan memastikan keluaran yang dihasilkan sesuai dengan yang diinginkan. Pengujian ini menggunakan perhitungan akurasi.

2.3 Fast Fourier Transform (FFT)

Fast Fourier Transform (FFT) adalah suatu transformasi yang mengubah data digital ke domain frekuensi. FFT merupakan salah satu algoritma yang paling sering digunakan dalam menganalisis dan manipulasi data digital. FFT sudah banyak digunakan untuk banyak hal, seperti *electroacoustic music* dan pengolahan sinyal audio, pengolahan citra, *medical imaging*, *pattern recognition*, *computational chemistry*, dan lain-lain [9]. Pada penelitian ini *Fast Fourier Transform (FFT)* digunakan sebagai ekstraksi fitur yang sebelumnya menggunakan nilai hasil dari perhitungan preemphasis dan windowing kemudian digunakan rumus *fast fourier transform* untuk mendapatkan nilai ekstraksi fitur suara.

Contoh perhitungan inidimiskalkan memiliki 2 jenis *file* suara untuk kata bola dan waktu. Masing-masing *file* suara memiliki 4 nilai sinyal suara. Untuk menghitung nilai ekstraksi ciri suara menggunakan *fft* maka digunakanlah persamaan (1) [9]. Langkah-langkah yang digunakan pada bagian ekstraksi ciri *Fast Fourier Transform (FFT)* adalah sebagai berikut:

- **Preemphasis.** Pada langkah ini, proses perubahan sinyal *speech*/sinyal ucapan menjadi sinyal rata. Sinyal suara $s(n)$ dimasukkan ke dalam sistem digital orde rendah yang digunakan untuk meratakan spectrum sinyal. Persamaan *preemphasis* yang paling umum digunakan ialah:

$$\tilde{s}(n) = s(n) - \tilde{a} \cdot s(n-1) \quad (1)$$

Dimana $\tilde{a}$ merupakan konstanta *filter preemphasis*, biasanya bernilai 0.9, maka 90% dari setiap sampel dianggap dari berasal dari sampel sebelumnya.

- **Windowing.** Pada tahap ini sinyal suara yang dipotong-potong menjadi beberapa *frame* akan dapat menyebabkan kesalahan data pada proses *Fourier transform*. *Windowing* digunakan untuk mengurangi *discontinuitas* sinyal pada awal dan akhir *frame*. Jika *window* didefinisikan sebagai $w(n)$, maka hasil dari penjedeleaan sinyal jelaskan pada persamaan (2).

$$y(n) = x(n)w(n), \quad \text{dimana} \quad 0 \leq n \leq N-1 \quad (2)$$

Pada penelitian ini *window* yang digunakan adalah *Hamming Window* yang dituliskan pada persamaan (3).

$$w(n) = 0,54 + 0,46 \cos\left(\frac{2\pi n}{N-1}\right), 0 \leq n \leq N-1 \quad (3)$$

- **Fast Fourier Transform (FFT)** adalah sebuah algoritma cepat untuk implementasi DFT yang dioperasikan pada sebuah sinyal waktu diskret yang terdiri dari N sampel, maka :

$$F(n) = \sum_{x=0}^{N-1} F(x) = e\left(\frac{-2\pi nx}{N}\right) \quad (4)$$

Berdasarkan persamaan *euler* yang menyatakan:

$$e^{\pm ix} = \cos(x) \pm i \sin(x) \quad (5)$$

Maka persamaan (4) dapat ditulis sebagai berikut:

$$F(x) = \sum_{x=0}^{N-1} \left[f(x) \left(\cos\left(\frac{2\pi nx}{N}\right) - i \sin\left(\frac{2\pi nx}{N}\right) \right) \right] \quad (6)$$

2.4 Bayes

Metode *Bayes* merupakan metode yang baik didalam mesin pembelajaran berdasarkan data training, dengan menggunakan probabilitas bersyarat sebagai dasarnya [4]. Definisi lainnya ialah *Bayes* dalam probabilitas dan statistika menunjukkan hubungan antara dua probabilitas kondisional dimana kedua kondisi tersebut saling bertolak belakang dan memperhitungkan bahwa probabilitas suatu kejadian (hipotesis) bergantung pada keadaan lain (bukti)[7]. Metode *Bayes* digunakan untuk menentukan hasil pengenalan dari beberapa data *template* latih yang telah disimpan sebelumnya. Untuk menghitung probabilitas *prior* untuk tiap *class* maka digunakan rumus sebagai berikut:

$$P(C) = \frac{N_j}{N} \quad (7)$$

Dimana:

$P(C)$ = probabilitas *prior* tiap *file*

N_j = jumlah file suara

N = jumlah total nilai suara

Menghitung *prior* tiap kelas, pada *class* C dengan rumus:

$$P(d|C) = P(x_1, x_2, \dots, x_n|C) \quad (8)$$

Dimana:

$P(d|C)$ = probabilitas suatu suara d didalam *file* C

C = *file*

d = suara

x = vector dari nilai atribut n

Hasil perhitungan $P(d|C)$ kemudian digunakan untuk menghitung probabilitas *posterior*.

$$P(C|d) = \frac{P(d|C)P(C)}{P(d)} \quad (9)$$

Dimana:

$P(C|d)$ = probabilitas suatu *file* C terhadap suatu suara d

$P(d|C)$ = probabilitas suatu suara d di dalam *file* C

$P(C)$ = probabilitas dari suatu *file* C

$P(d)$ = probabilitas dari suatu suara

Pendekatan *Bayes* akan memilih kategori yang memiliki probabilitas paling tinggi (*Cmap*) yaitu:

$$C_{map} = \operatorname{argmax} \frac{P(C)P(d|C)}{P(d)} \quad (10)$$

Nilai $P(x)$ dapat diabaikan karena nilainya konstan untuk semua C, dapat dituliskan sebagai berikut:

$$C_{map} = \operatorname{argmax} P(C)P(d|C) \quad (11)$$

Teorema *Bayes* yang digunakan pada proses pengambilan keputusan tidak terlepas dari teori peluang sebagai konsep dasar. Teorema *Bayes* dikenal sebagai rumus dasar untuk peluang bersyarat yang tidak bebas.

3. HASIL DAN PEMBAHASAN

Bagian ini akan menjelaskan hasil yang dilakukan pada penelitian dan pembahasannya sebagai berikut:

3.1 Hasil

Pada penelitian ini ruang lingkup yang digunakan dibagi menjadi dua yaitu kebutuhan non fungsional dan kebutuhan fungsional

3.1.1 Kebutuhan Non Fungsional

Kebutuhan non fungsional menggambarkan kebutuhan sistem yang menitik beratkan pada properti perilaku yang dimiliki sistem, diantaranya kebutuhan perangkat lunak (*Software*).

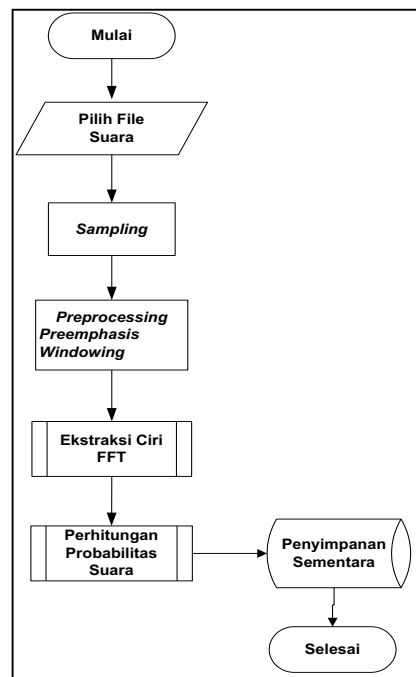
- Kebutuhan Perangkat Lunak (*software*)

Perangkat lunak yang dibutuhkan dalam pembuatan sistem pengenalan pola suara manusia berekstensi *file wav* menggunakan metode *Fast Fourier Transform* dan *Bayes* yaitu:

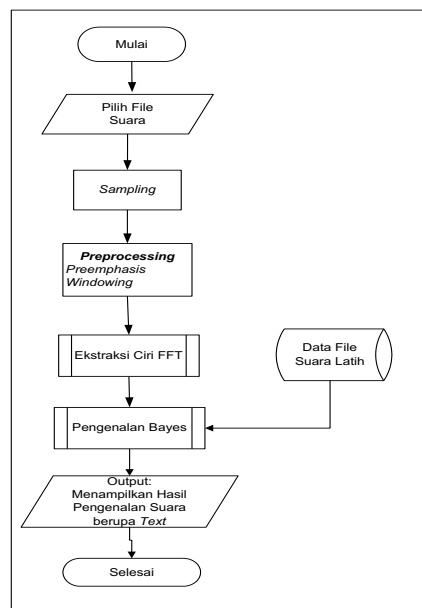
1. *Sistem operasi Windows 8*
2. *Browser* : *Mozilla Firefox* dan *Google Chrome*
3. *Matlab R2016a* untuk bahasa pemrograman
4. *Microsoft Visio 2007* dan *Astah Community 6.6.3 (ModelVersion : 36)* untuk perancangan.

3.1.2 Kebutuhan Fungsional

Kebutuhan fungsional membahas tentang perancangan sistem di dalam sistem yang akan dibangun. Kebutuhan fungsional pada sistem ini menggunakan *flowchart*. *Flowchart* merupakan bagian dari sistem yang mendeskripsikan alur proses sistem yang akan dibangun. Alur proses keseluruhan sistem pengenalan pola suara manusia berekstensi *file wav* menggunakan metode FFT dan *Bayes*. Dimana proses pertama dimulai dengan melatih data suara terlebih dahulu dengan cara *user* memilih *file* suara yang telah tersedia kemudian masuk ke proses *sampling* dan dilanjutkan ke tahap *preprocessing* yang hasilnya akan dilakukan untuk pengekstrasian ciri menggunakan *Fast Fourier Transform* (FFT). Data latih akan melakukan proses yang sama dengan data uji, hanya saja data latih akan dijadikan sebagai template untuk menghitung nilai probabilitas suara antara data latih dan data uji, sehingga data latih akan dilakukan terlebih dahulu. Sedangkan masukan dari data uji hasilnya akan langsung dibandingkan dengan data latih menggunakan metode *Bayes*. *Bayes* digunakan untuk menentukan hasil pengenalan dari beberapa data *template* latih yang telah disimpan sebelumnya. Proses pengenalan suara akan mengeluarkan hasil berupa *text* yang akan tertera pada layar antarmuka. Didalam proses *preprocessing* terdapat subproses yaitu *preemphasis* dan *windowing*. Pada gambar 1 terdapat *flowchart* alur proses data latih dan gambar 2 terdapat *flowchart* alur proses data uji sistem pengenalan pola suara.



Gambar 1 Alur Proses Data Latih



Gambar 2 Alur Proses Data Uji

3.2 Pembahasan

Pada penelitian ini pengujian dilakukan oleh 4 (empat) orang responden yang terdiri dari 2 (dua) orang laki-laki dan 2 (dua) orang perempuan. Setiap orang akan mengucapkan 1 (satu) kata sebanyak 3 kali. Untuk pengujian suara laki-laki 1 dan perempuan 1 diambil dari pola referensi suara yang ada didalam *database* atau suara yang telah terlatih. Sedangkan untuk pengujian suara laki-laki 2 dan perempuan 2 tidak mempunyai pola referensi suara didalam *database* atau disebut pola suara baru yang belum dilatih. Masing-masing perekaman berdurasi lebih kurang 2 sampai 3 detik dan memiliki kapasitas rata-rata sebesar 80KB yang berbentuk *file wav*.

3.2.1 Algoritma Fast Fourier Transform

Untuk menghitung nilai ekstraksi ciri suara menggunakan algoritma FFT maka digunakanlah persamaan (6) sebagai berikut:

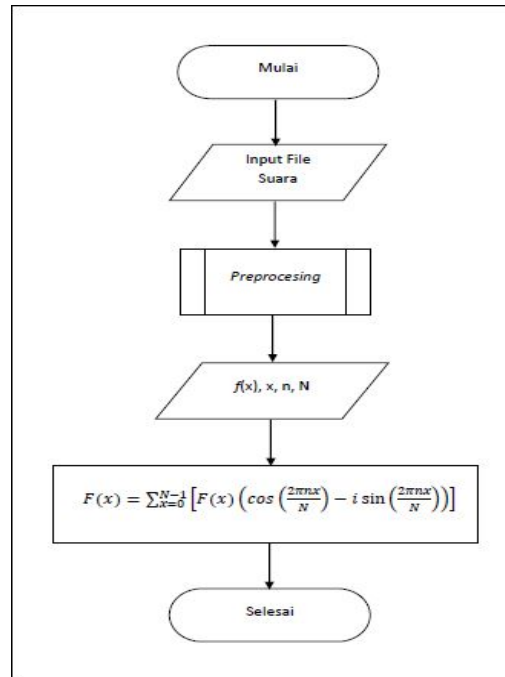
$$F(x) = \sum_{n=0}^{N-1} \left[f(x) \left(\cos \left(\frac{2\pi nx}{N} \right) - i \sin \left(\frac{2\pi nx}{N} \right) \right) \right] \quad (12)$$

Dimana:

$F(x)$ = Keluaran dalam domain frekuensi

x = indeks dalam domain waktu (0,1,..., N-1)

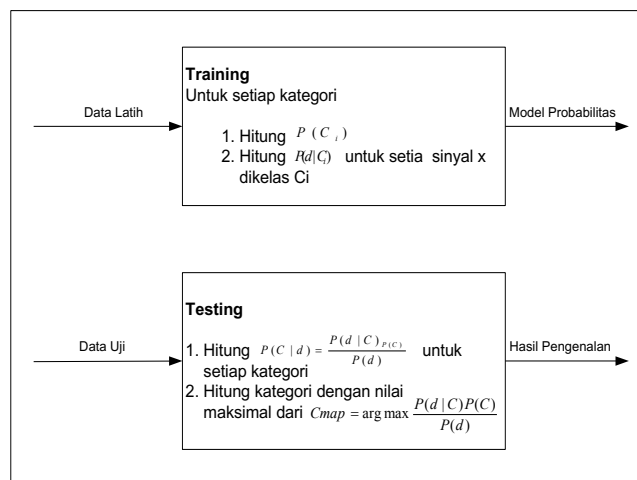
n = indeks dalam domain frekuensi (0,1,..., N-1)



Gambar 3 Flowchart Algoritma FFT

3.2.2 Metode Bayes

Pada perhitungan metode *bayes* digunakan untuk proses pengenalan suara. Metode ini terdiri dari 2 (dua) proses yaitu proses *training* dan proses *testing*.



Gambar 4 Flowchart Algoritma Bayes

- Proses *training* metode *Bayes*

Hal yang dilakukan pada proses *training* metode *Bayes* yaitu untuk menghitung nilai probabilitas setiap sinyal suara yang terucap agar dapat dikenali.

Untuk menghitung probabilitas prior untuk tiap file maka digunakanlah persamaan (7) sebagai berikut:

$$P(C) = \frac{N_j}{N} \quad (13)$$

Setelah menghitung prior tiap *file*, selanjutnya hitunglah setiap sinyal yang terdapat pada *file* C dengan persamaan menggunakan persamaan (8) sebagai berikut:

$$P(d|C) = (x_1, 2, \dots, x_n | C) \quad (14)$$

- Proses *testing* metode *Bayes*

Pada tahap ini, data uji akan melewati proses pengenalan berdasarkan data latih pada proses *training*. Hal pertama yang dilakukan yaitu menghitung probabilitas *posterior* dengan persamaan (9)

$$P(C|d) = \frac{P(d|C)P(C)}{P(d)} \quad (15)$$

Nilai $P(x)$ dapat diabaikan karena nilainya konstan untuk semua C , maka digunakanlah persamaan (11) sebagai berikut:

$$C_{map} = \operatorname{argmax} P(C)P(d|C) \quad (16)$$

Setelah melakukan perhitungan pada tiap file maka di dapatlah file suara yang dikenali adalah bola berdasarkan nilai terbesar pada perhitungan yang telah dilakukan.

3.2.3 Implementasi dan Pengujian

Bagian ini akan menggambarkan implementasi dan pengujian yang dilakukan pada penelitian sebagai berikut:

3.2.3.1 Implementasi

Hasil implementasi pada aplikasi ini dimulai dengan memilih *file* suara yang terdapat dalam folder suara yang akan diuji. Setelah *file* suara dipilih maka akan tampil bentuk sinyal *suarasampling*, *preemphasis*, *windowing* dan ekstraksi dari FFT.

- **Tampilan Halaman Menu Utama**

Pada menu utama ini terdapat 3 tombol *button* yaitu tombol pengenalan pola suara untuk menguji coba suara, tombol tentang untuk melihat informasi tentang program yang dibangun, dan tombol keluar untuk keluar dari program.

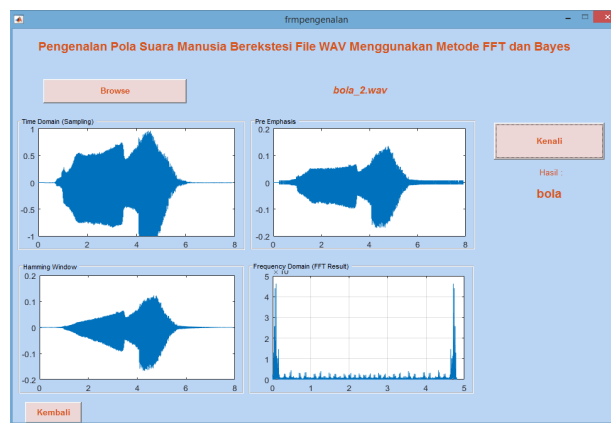


Gambar 4 Tampilan Antar Muka Gambar

• Tampilan Halaman Pengenalan Pola Suara

Halaman pengenalan pola suara ialah halaman yang digunakan untuk menguji dan mengenali suara, file suara yang didapat melalui proses perekaman sesuai dengan klasifikasi yang telah ditetapkan, pada sistem pengenalan suara yang dibuat ini data uji yang digunakan sebanyak 120 data suara yang terdiri dari 4 orang responden 2 laki-laki dan 2 perempuan. Kata-kata yang disimpan didalam database ialah bola, dadu, fakta, foto, guru, harta, lampu, mata, salju, waktu yang masing-masing diucapkan sebanyak 3(tiga) kali. Pada gambar 4 dapat dilihat tampilan antar muka pengenalan pola suara yang terdiri dari 3 (tiga) tombol *button* yaitu *Browse*, *Kenali* dan *Kembali*. Terdapat 4 (empat) *axes* yaitu: *sampling*, *preemphasis*, *windowing*, ekstraksi FFT dan terdapat 1 (satu) label untuk keluaran hasil yang dikenali.

Untuk menguji file suara *user* dapat menekan tombol *browse* terlebih dahulu untuk memilih *file* suara yang terdapat dalam folder yang akan diuji. Setelah *file* suara dipilih maka akan tampil bentuk sinyal suarasampling, *preemphasis*, *windowing* dan ekstraksi ciri FFT. Setelah bentuk sinyal suara tampil untuk mengenali suara apa yang diuji coba maka *user* dapat menekan tombol *kenali*, setelah ditekan tombol *kenali* maka hasil dari suara tersebut akan keluar berupa *text*.



Gambar 5 Tampilan Pengenalan Pola Suara

3.2.3.2 Pengujian

Pengujian ini dilakukan untuk mengetahui tingkat ketepatan sistem pada pengenalan pola suara. Pada pengujian ini dilakukan oleh 4 orang responden yaitu 2 laki-laki dan 2 perempuan. Untuk pengujian suara yang di ambil dari 2 orang responden yang datanya telah tersimpan didatabase dinamai laki-laki 1 dan perempuan 1. Sedangkan untuk data suara yang belum tersimpan didatabase dinamai laki-laki 2 dan perempuan 2.

Tabel 1 Hasil pengujian suara laki-laki 1

No	Nama File	Dikenali Sebagai	Keterangan
1	bola_1.wav	Bola	Benar
2	bola_2.wav	Bola	Benar
3	bola_3.wav	Bola	Benar
4	dadu_1.wav	Dadu	Benar
5	dadu_2.wav	Dadu	Benar
6	dadu_3.wav	Dadu	Benar
7	fakta_1.wav	Fakta	Benar
8	fakta_2.wav	Fakta	Benar
9	fakta_3.wav	Fakta	Benar
10	foto_1.wav	Foto	Benar

No	Nama File	Dikenali Sebagai	Keterangan
11	foto_2.wav	Foto	Benar
12	foto_3.wav	Foto	Benar
13	guru_1.wav	Guru	Benar
14	guru_2.wav	Guru	Benar
15	guru_3.wav	Guru	Benar
16	harta_1.wav	Harta	Benar
17	harta_2.wav	Harta	Benar
18	harta_3.wav	Harta	Benar
19	lampu_1.wav	Lampu	Benar
20	lampu_2.wav	Lampu	Benar
21	lampu_3.wav	Lampu	Benar
22	mata_1.wav	Foto	Salah
23	mata_2.wav	Mata	Benar
24	mata_3.wav	Mata	Benar
25	salju_1.wav	Salju	Benar
26	salju_2.wav	Salju	Benar
27	salju_3.wav	Salju	Benar
28	waktu_1.wav	Waktu	Benar
29	waktu_2.wav	Waktu	Benar
30	waktu_3.wav	Waktu	Benar

Keterangan: Benar (Suara yang dikenali). **Salah** (Suara yang tidak dikenali)

Dari tabel 1 didapatkan:

Jumlah data uji salah = 1

Jumlah data uji benar = 29

Perhitungan akurasi pada tabel 1 ialah sebagai berikut:

$$Accuracy = \frac{29}{30} \times 100\% = 96\%$$

Jadi didapatkan hasil akurasi dari suara laki-laki 1 sebesar 96% hal ini dapat dikatakan bahwa pengujian untuk suara laki-laki 1 memberikan hasil yang baik.

Tabel 2 Hasil pengujian suara perempuan 2

No	Nama File	Dikenali Sebagai	Keterangan
1	bola_1.wav	Mata	Salah
2	bola_2.wav	Fakta	Salah
3	bola_3.wav	Fakta	Salah
4	dadu_1.wav	Mata	Salah
5	dadu_2.wav	Dadu	Benar
6	dadu_3.wav	Mata	Salah
7	fakta_1.wav	Fakta	Benar
8	fakta_2.wav	Fakta	Benar
9	fakta_3.wav	Fakta	Benar
10	foto_1.wav	Waktu	Salah
11	foto_2.wav	Mata	Salah
12	foto_3.wav	Foto	Benar
13	guru_1.wav	Salju	Salah
14	guru_2.wav	Guru	Benar

No	Nama File	Dikenali Sebagai	Keterangan
15	guru_3.wav	Mata	Salah
16	harta_1.wav	Fakta	Salah
17	harta_2.wav	Mata	Salah
18	harta_3.wav	Mata	Salah
19	lampu_1.wav	Mata	Salah
20	lampu_2.wav	Lampu	Benar
21	lampu_3.wav	Lampu	Benar
22	mata_1.wav	Mata	Benar
23	mata_2.wav	Mata	Benar
24	mata_3.wav	Mata	Benar
25	salju_1.wav	Salju	Benar
26	salju_2.wav	Mata	Salah
27	salju_3.wav	Salju	Benar
28	waktu_1.wav	Waktu	Benar
29	waktu_2.wav	Mata	Salah
30	waktu_3.wav	Waktu	Benar

Keterangan: Benar (Suara yang dikenali). **Salah** (Suara yang tidak dikenali)

Dari tabel 2 didapatkan:

Jumlah data uji salah = 15

Jumlah data uji benar = 15

Perhitungan akurasi untuk suara perempuan 2 ialah sebagai berikut:

$$Accuracy = \frac{15}{30} \times 100\% = 50\%$$

Jadi didapatkan hasil akurasi dari suara perempuan 2 sebesar 46% hal ini dapat dikatakan bahwa pengujian untuk suara perempuan 2 memberikan hasil yang kurang baik.

Untuk tabel pengujian perempuan 1 dan laki-laki 2 sama seperti tabel pengujian laki-laki 1 dan perempuan 2. Hanya saja memiliki memiliki hasil akurasi yang berbeda-beda.

Hasil pengujian untuk suara perempuan 1 didapatkan:

Jumlah data uji salah = 7

Jumlah data uji benar = 23

Perhitungan akurasi untuk suara perempuan 1 ialah sebagai berikut:

$$Accuracy = \frac{23}{30} \times 100\% = 76\%$$

Jadi didapatkan hasil akurasi dari suara perempuan 1 sebesar 76% hal ini dapat dikatakan bahwa pengujian untuk suara perempuan 1 memberikan hasil baik.

Hasil pengujian untuk suara laki-laki 2 didapatkan:

Jumlah data uji salah = 16

Jumlah data uji benar = 14

Perhitungan akurasi untuk suara laki-laki 2 ialah sebagai berikut:

$$Accuracy = \frac{14}{30} \times 100\% = 46\%$$

Jadi didapatkan hasil akurasi dari suara laki-laki 2 sebesar 46% hal ini dapat dikatakan bahwa pengujian untuk suara laki-laki 2 memberikan hasil yang kurang baik.

4. KESIMPULAN

Berdasarkan hasil penelitian tentang pengenalan pola suara manusia berekstensi file wav menggunakan metode *Fast Fourier Transform* dan *Bayes*, maka dapat ditarik kesimpulan bahwa pengenalan pola suara yang dibangun ini mampu mengenali suara yang menghasilkan *text* sesuai dengan kata yang diucapkan dengan baik untuk suara orang yang tersimpan dalam *database* dengan tingkatakurasi untuk suara laki-laki pertama sebesar 96% dan suara perempuan pertama sebesar 76%. Sedangkan kelemahan sistem adalah pengenalan suara belum dapat dikenali dengan baik dengan data uji yang an tidak tersimpan dalam *database*, dimana tingkat akurasi yang didapat untuk pengujian suara laki-laki kedua sebesar 46% dan perempuan kedua sebesar 50%.

UCAPAN TERIMA KASIH

Terimakasih kepada Dosen Pembimbing saya Rita Wiryasaputra, ST.,M.Cs dan Lastri Widya Astuti, M.Kom yang selalu meluangkan waktu untuk membimbing saya dalam menyelesaikan artikel ilmiah ini.

DAFTAR PUSTAKA

- [1] A. Qustoniah., 2012, Perancangan Sistem Pengenalan Pola Suara Menggunakan Metode Piecewise Dynamic Time Warping, *Widya Teknika*, No. 1, Vol. 20, hal. 11-17.
- [2] E. Ronando., dan M. Isa Irawan., 2012, Pengenalan Ucapan Kata Sebagai Pengendali Metode Linear Predictive Coding – Neuro Fuzzy, *J. Sains dan Seni ITS*, No. 1, Vol. 1, hal 51-56.
- [3] M. T. S. Jaya., D. Puspitaningrum., dan B. Susilo., 2016, Penerapan Speech Recognition Pada Permainan Teka-Teki Silang Menggunakan Metode Hidden Markov Model (HMM) Berbasis Desktop, *Jurnal Rekrusif*, No. 1, Vol. 4, hal 119-129.
- [4] S. Rahayu., 2013, Sistem Pakar Untuk Mendiagnosa Penyakit Gagal Gijal Dengan Menggunakan Metode Bayes, *Pelita Infromatika Budi Darma*, No. 3, Vol. 6, hal. 129–134.
- [5] N. Fetra. Dan M. Irsyad., 2015, Aplikasi Pencarian Chord dalam Membantu Penciptaan Lagu Menggunakan Algoritma Fast Fourier Transform (FFT) dan Metode Klasifikasi K-Nearest Neighbor (KNN), *Jurnal CoreIT*, No. 2 Vol. 1, hal. 30-36.
- [6] F. Liantoni., dan H. Nugroho., 2015, Klasifikasi Daun Herbal Menggunakan Metode Naïve Bayes Classifier Dan K- Nearest Neighbor, *Jurnal SimanteC*, No. 1, Vol. 5, hal. 9-16.
- [7] L. Wahyuni., dan S. Darma, 2015, Implementasi Teorema Bayes Dalam Menentukan Varietas, *Seminar Nasional Teknologi Informasi dan Multimedia*, hal. 6–8.
- [8] A. Rosa and S. M., 2013, *Rekayasa Perangkat Lunak*, Informatika Bandung, Bandung.
- [9] H. Harahap., G. Budiman., dan L. Novamizanti., 2016, Implementasi Teknik Watermarking menggunakan FFT dan Spread Spectrum Watermark pada Data Audio Digital, *Jurnal Elkomika*, No. 1, Vol. 4, hal. 2459–9638.